Regional convergence across European Union

Cristina Brasili*, Luciano Gutierrez‡

*Department of Statistics, University of Bologna, Italy
‡Department of Agricultural Economics, University of Sassari, Italy

February, 2004

Please, do not quote without permission

Abstract

This paper analyzes the per-capita incomes convergence process across 140 NUTS2 European regions during the period 1980-1999. Two methods of analysis have been used. The first adopts the non parametric method proposed by Quah (1996, 1997) to study whether the cross-regions income distribution shows evidence of convergence, i.e a tendency for the steady-state distribution to cluster around one or more poles of attraction, or divergence. The second uses panel unit root tests for cross-sectionally correlated panels. Unlike other studies, we find evidence of convergence among the EU regions. When looking at the distribution of per-capita income, we show that it converges toward the average pole, without convergence clubs emerging. Panel unit root tests strongly reject the null of divergence for the full sample of regions and evidence is also provided for two regional subgroups.

Key words **:** Convergence, Distribution analysis, Panel unit root tests.

JEL classification: C22, C23, O41, O47

Correspondence :
Luciano Gutierrez
Department of Agricultural Economics
University of Sassari
Via E. De Nicola 1, Sassari 07100
Italy

Tel.: +39.079.229.256
Fax: +39.079.229.356
e-mail: lgutierr@uniss.it
web: http://www.gutierrezluciano.net

1. Introduction

This paper focuses on the question of per-capita income convergence inside the European Union (EU) regions. The convergence or divergence issues have been much debated in recent years. As is well known, the standard neoclassical growth model (Solow, 1956; Swan, 1956) asserts that per-capita output across countries or regions converges when they have similar preferences, technology levels and institutional and legal systems. Thus gaps in national or regional outputs must disappear over time. On the other hand, the endogenous growth model (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992) asserts that per-capita income is mainly influenced by country-specific factors which endogenously influence output dynamics. If this is the case, countries will not converge over time given that per-capita income only responds to country specific factors.

It is useful to know if EU integration, that is the elimination of trade barriers and adoption of common trade, industrial, fiscal and monetary policies, has spurred regional convergence. This question assumes greater importance when related to the next EU enlargement, since the new member countries have per-capita incomes well below the current EU average, and the implications of convergence, or lack of convergence, are likely to have a substantial impact on the EU's regional cohesion.

Many studies have been presented in recent years in an attempt to clarify this question. Barro and Sala-I-Martin (1995), focusing on per-capita incomes for 90 EU regions during the period 1950-1990 and using cross-section regressions analysis, conclude that there are signs of conditional convergence across the EU regions, that is regions converge towards different levels of per-capita income but regions that are further from their steady-state per-capita income level will grow faster. Boldrin and Canova (2001), using a similar methodology, severely criticised the previous results. Using a different dataset, which includes 185 NUTS2 EU regions observed during the period 1980-1996, they concluded that the results are mixed and not supportive of convergence of regional per-capita income. Canova and Marcet (1995) also, basing the analysis on per-capita incomes for 144 NUTS2 EU regions, found only limited signals of convergence during the period 1980-1992. Finally Canova (2004), adopting a new methodology to analyse the distribution density of per-capita incomes and the same dataset as that used by Canova and Marcet (1995), concluded that the steady-state distribution tended to cluster around four poles of attractions characterised by different dynamics, different steady-states and different mobility features.

During recent years the empirical literature on income convergence has evolved. As previously noted, the early studies were mainly based on cross-country analysis, which regress the average per-capita income growth rates on initial income level. Negative correlation between income growth and initial income is interpreted as evidence of the convergence hypothesis. The appropriateness of the cross-country regression method has been criticized by Quah (1993a) who shows that an inverse relationship between income growth and initial income is consistent with a stable variance in cross-country variance. Bernard and Durlauf (1996) highlight that cross-section tests tend to spuriously reject the null of convergence when countries have different long run steady states, or in other words, transition matters in determining income dynamics.

Two different methodologies have been employed to resolve the previous problems. The first procedure was proposed and applied in a number of papers by Quah (1993b, 1996) where the entire distribution was studied to assess cross-countries convergence of per-capita incomes. These studies usually highlight the formation of “twin peaks” or convergence clubs, i.e the polarization of the income distribution into “peaks” or “clubs” of rich and poor countries. The second methodology uses panel unit root and cointegration tests to evaluate per-capita income convergence (Evans, 1998; Phillips and Sul 2003; Cheung and Pascual, 2004) which are better suited for taking into account income heterogeneity across countries or regions and over time.

In the next section we briefly review both methodologies while in the third section we apply both procedures to the per-capita incomes of a large sample of 140 EU regions observed during the period 1980-1999. In brief we find that, contrary to part of the previous literature, poorer EU regions are catching up with richer regions. These findings, although we emphasize that further research is needed, may indicate first that EU integration has enhanced per-capita income convergence and second that EU newcomers may find new tools to reduce income inequalities with respect to older EU regions.

2. A briefly review of distribution dynamics and panel analysis

Before introducing more sophisticated methodologies, it is useful to look at the evolution of European per-capita income trends during the last twenty years.[1] In Table 1 we present the trend coefficients of the log of per-capita income for 140 EU regions during the period 1980-1999.

Table 1 about here

The 140 regions are ranked according to their 1980 per-capita income. Thus the poorest region in 1980 is labeled 1 and the wealthiest region is labeled 140. During the period of analysis the EU regions experienced a significantly positive trend coefficient of 0.057. When dividing the sample in half, the halves exhibit different trends. The poorer half shows a higher trend coefficient, 0.060,

[1] See section 3 for a description of the dataset.

than the wealthier countries, 0.053. When the original sample was divided into three, four, five or more groups one fact emerged clearly: relatively poorer regions show a higher, significant and positive trend coefficient.

From the previous analysis it seems that in the last twenty years per-capita incomes in the EU regions have shown significant convergence patterns, in the sense that the poorest regions grew more than richest. However this conclusion may be biased for the following reasons. The results can be sensitive to the inclusion or exclusion of a region from a group. If only one region in the wealthiest group has a relevant negative trend during the period, this may bias the conclusion towards finding convergence. Thus studying the entire per-capita income distribution over time as in Quah (1993b, 1996) can protect one from potentially erroneous conclusions and shed light on the convergence process and the appearance of convergence clubs. Secondly estimating a trend-stationary process can have serious pitfalls if the variables are non stationary. If this is the case, as it is for many economics variables, models based on difference-stationary process or unit root processes are more appropriate. Both methodologies will be presented in the following sections.

2.1 Distribution dynamics

As previously stated, the purpose of distribution dynamics is to study how the cross-section distribution of certain economic variables, in our case per-capita income across European regions, evolves over time. Because the methodology has been presented in a number of papers (see for example Quah, 1993a, 1996, 1997), here we will give only a brief sketch. The reader can refer to these works for a deeper analysis.

Assume that the per-capita income y_t can take values inside a certain finite set E. The distribution of that variable at time *t*, labeled F_t, is time-invariant. Define f_t the associated probability measures of F_t. The dynamics of f_t can be modeled as a first order autoregressive process

$$f_t = M' f_{t-1},\ t \geq 1 \tag{1}$$

When y_t is discrete, the matrix *M* is usually defined as the transition probability of a Markov process, i.e. each element in *M* describes the probability of transition from a given state to another state in one step. However if y_t can take infinite values, i.e E is an uncountable set, we need a continuous counterpart for *M*. Let A be a subset of E and define a new function $\mathrm{P}(y, A)$, called stochastic transition function or stochastic kernel. This function describes the conditioned probability that in the next period the per-capita income will take a value in the set A, given that in the previous period it is in the state y, i.e $\mathrm{P}(y,A)=\Pr(y_t \in A | y_{t-1}=y)$. Thus the per-capita income distributions in the two periods will be linked by the following relationship

$$F_t = \int \mathrm{P}(y,A) F_{t-1} dy \tag{2}$$

In the empirical section an estimate of $\mathrm{P}(y, A)$ for the EU regions will be presented.

2.2 Panel analysis

In order to analyze the panel unit root approach to convergence let us start by introducing the following panel regression equation

$$\ln\left(\frac{y_{it}}{y_{\cdot t}}\right) = a_i + r_i \ln\left(\frac{y_{it-1}}{y_{\cdot t-1}}\right) + \sum_{s=1}^{p_i} b_{is}(1-L)\ln\left(\frac{y_{it-s}}{y_{\cdot t-s}}\right) + u_{it} \tag{3}$$

where y_{it} is the observed per-capita income for region i at time t, $y_{\cdot t}$ is the cross-section average in each period t, a_i, r_i, b_{ij} are parameters, L is the lag operator and u_{it} is a cross-sectional correlated error.[2] Looking at (3) one can easily infer that when $r_i = 1$ for all i, all regions diverge, thus not rejecting the null hypothesis of a common panel unit root is the same as accepting the hypothesis that during the period of analysis all regions do not convergence to the cross section average. The alternative hypothesis is usually stated as $H_1: r_i < 1$, at least for one region. This means that rejecting the null hypothesis does not imply overall convergence. Some regions, or better some subgroups of regions, may converge and others may not. Thus some method to subgroup the regions is needed and this will be supplied in the empirical section.

Before introducing results it is useful to answer the following questions. Why do we focalize the attention on panel regression and not on time series regression and why do we introduce a degree of cross-sectional dependence in (3)?

Over the last few years, a great deal of attention has been paid to the non-stationary property of panels. Starting from the seminal works of Quah (1990, 1994), Breitung and Meyer (1991) Levin and Lin (1992, 1993), and Im *et al.* (1997), many tests have been proposed which attempt to introduce unit root tests in panel data. These show that, by combining the time series information with that of the cross-section, the inference that unit roots exist can be more straightforward and precise, especially when the time series dimension of the data is relatively short, and similar data

[2] Regression (1) has been studied by Evans (1998). Bernard and Durlauf (1995) examined a similar regression where the cross-section average $y_{\cdot t}$ was substituted by the per-capita income of a reference country *j*.

may be obtained across a cross-section of units such as countries or commodities. In synthesis, panel unit root tests have higher power than time series unit root tests.

However all the panel unit root tests suffer from serious limitations when the cross-sectional units are correlated (see O'Connell, 1998). Some papers have been presented in recent years that address this issue. For example, Bai and Ng (2003), Moon and Perron (2003) and Phillips and Sul (2003) and Choi (2002) use common factor components. In brief, all the above mentioned works propose a factor model in which cross-sectional dependence is generated by one or more factors[3] which are common to all the individual units (but which may exert different effects on the individual unit) and by uncorrelated idiosyncratic shocks across all the individual units.

The cross-sectional dependence is modeled as

$$u_{it} = l_i^{'} f_t + e_{it} \tag{2}$$

where f_t are K vectors of unobservable factors, $l_i^{'}$ factor loading coefficient vectors and e_{it} are idiosyncratic shocks. Note that the panel unit root tests proposed by Levin and Lee (1992, 1993) and Im and al. (1997) fail to take account of cross-sectional dependence, causing on one hand huge size distortion of the tests and, on the other, introducing restrictive economic specification when, as seems to be the case especially for EU regions, per-capita incomes show strong cross-sectional correlation.

3. The data set and results

The data set used covers 140 NUTS2 European regions. The annual per-capita GDP in Purchasing Power Standard (PPS) units come from the Eurostat-REGIO dataset. They cover the period 1980-1999 and are only available for eleven countries. Regional data for Austria, Finland, Sweden and Ireland are not available and in these cases the aggregate value has been used. [4]

The estimated stochastic kernel and contour plots are presented in figures 1 and 2, describing ten-year-horizon evolutions of the distribution of per-capita income in each region relative to the EU average[5]. Figure 1 does not exhibit "twin peaks" as the cross-countries estimates usually do.

[3] This is not true for the Phillips and Sul (2003) tests where only one factor is permitted.
[4] The countries and the number of regions (in parenthesis) in the dataset are: Belgium (7); Denmark (1); Ireland(1); Luxemburg (1); Germany (27); Greece (13); Spain (18); France (21); Italy (20); Netherlands (12); Portugal (5); Austria (1); Finland (1); Sweden (1); United Kingdom (11).
[5] An Epanechnikov kernel estimator and the method proposed in Silverman (1986, 4.3.2) to calculate window width have been used. All these computations, and the following panel unit root tests, are performed using GAUSS 3.21 and the routines are freely available upon request from the authors.

Figure 1 and 2 about here

The contour plot of the surface shows that the values of relative per capita income less than 1.0 and greater than 1.0 lie respectively below and above the 45° line. This picture means that poorer regions tend to be more likely to increase their relative income over the ten-years horizon. The opposite is true for richer regions. Thus low per-capita income regions tend to grow more quickly than wealthier regions. Note that from the contour plot a "peak" for richer regions seems to emerge. However looking at figure 3, which plots the estimated ergodic (i.e. long-run) density distribution, we note that the density is strongly unimodal with a mean close to unity. Thus it seems from these figures that there has been a tendency to convergence across the European regions during the period of analysis.

Figure 3 about here

Additional information on the convergence process highlighted by distribution analysis can be obtained from looking at the results of panel data analysis. The first task, when computing multifactor analysis as in (3), is to specify the number of factors r correctly. We follow Bai and Ng (2002) and we use what they label BIC_3. These criteria defines the correct number of factors, taking into account the mean squared sum of residuals, plus a penalty function for over-fitting. Bai and Ng (2002) show that these criteria perform well for our size of data sample. We compute the number of factors using a maximum of 3. The BIC_3 criteria suggest that there are two common factors.

Before using panel unit root tests it is useful to analyze their size and power for our sample of data. We perform a Monte Carlo analysis, as in Gutierrez (2003), for a panel of 20 observations and 140 units. The results of the simulation, not reported for brevity, highlight that Moon and Perron's (2003) t_b statistic has correct size, 0.057, and high power, 0.995. Choi (2002) tests are strongly oversized while Bai and Ng (2003) and Phillips and Sul (2003) tests are downsized with low power.

Given the better properties of Moon and Perron's (2003) t_b test, only this statistic is reported in Table 2.

Table 2 about here

Looking at the result, the statistic rejects the null of non-stationarity when the full sample of regions are analyzed, i.e. the test statistic rejects the hypothesis of divergence across the regions during the period of analysis. Naturally this does not mean that all regions converge, indeed some can actually diverge and some converge. To analyze this we split the sample of regions into different groups. The first group shows possible differences in the convergence process, by splitting the sample into all the regions that we call "peripheral" regions [6] and the remaining regions which can be defined "core" regions. Interestingly, we note first that *t_b* test statistic rejects the null, indicating that per capita incomes in the two samples of regions converge toward the EU average and second, looking at p-values, the peripheral regions show a more intensive convergence process. A similar picture emerges when splitting the sample of regions into the "old" regions and "newcomers" that entered the EU after 1980 [7]. In both cases the test statistic rejects the null of per-capita income divergence. Note also that in this case the p-values indicate stronger convergence for the "newcomer" regions.

4. Conclusions

A large amount of literature on convergence across economies has produced a great deal of both non parametric and parametric methodologies. Basically, non parametric methodologies study the distribution dynamics of cross-section per-capita incomes. This process could have many limiting outcomes, from complete equality (convergence) to increasing inequality (divergence) or polarization around peaks. The second group of parametric methodologies adopts cross-section, time series analyses and more recently panel data approaches which permit cross-section correlation across the economies to be taken into account.

We use both methodologies, distribution dynamics and panel data analyses, to study the growth convergence process of per-capita incomes for a group of 140 European regions during the period 1980-1999. We find considerable evidence of convergence. Distribution analysis highlights that per-capita income of poorer European regions converge toward the mean, and no evidence of polarization into "twin peaks" is found. Panel data analysis confirms these findings. The convergence process is more intense for low income regions.

In conclusion, the results indicate first that EU integration may have enhanced per-capita income convergence processes and second that the peripheral regions, or the regions that entered in EU after 1980, have converged more rapidly toward the average per-capita EU income. These findings may be good news for EU newcomers. They may be able to find new tools to reduce income inequalities with respect to older EU regions more rapidly.

[6] We include in the group of "peripheral" Greece, Ireland, Portugal, Spain, South of Italy and North of UK regions.
[7] We insert in this group Portugal, Spain, Finland and Sweden regions.

References

Aghion, P., and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2) : 323-351.

Bai, J., Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 70:1, 191-221.

Bai, J., Ng, S. (2003). A PANIC Attack on Unit Roots and Cointegration, mimeo, Boston College.

Barro, R.J., and Sala-i-Martin, X. (1995). *Economic Growth*. New-York: McGraw-Hill.

Bernard, A.B., Durlauf, S.N. (1995). Convergence in international output. *Journal of Applied Econometrics*, 10, 97-108.

Bernard, A.B., Durlauf, S.N. (1996). Interpreting test of convergence hypothesis. *Journal of Econometrics*, 71, 161-174.

Breitung, J., Meyer, W. (1991). Testing for Unit Roots in Panel Data: are Wages on Different Bargaining Levels Cointegrated? Institute für Wirtschaftsforschung Working Paper, June.

Boldrin M., Canova F. (2001). Inequality and Convergence in Europe's regions: reconsidering European regional policies, Economic Policy, 32, 207-245.

Canova, F. (2004). Testing for convergence clubs in income per capita: a predictive density approach, *International Economic Review*, 45, 49-77.

Canova, F., Marcet A. (1995). The poor stay poor: non convergences across countries and regions, CEPR working paper n.° 1405.

Cheung, Y., Pascual A.G. (2004), Testing output convergence: a re-examination. Oxford Economic Papers, 56, 45-63.

Choi, I. (2001). Unit Root Tests for Panel Data. *Journal of International Money and Finance*, 20, 249-272.

Choi, I. (2002). Combination Unit Root Tests for Cross-Sectionally Correlated Panels, mimeo, Hong Kong University of Science and Technology.

Elliot, G., Rothemberg T.J. and Stock J.H. (1996), Efficient Tests for an Autoregressive Unit Root, *Econometrica*, 64, 813-836.

Evans P. (1998) Using panel data to evaluate growth theories. *International Economic Review*, 39, 295-306.

Grossman, G. and Helpman, E. (1991). *Innovation and growth in the global economy*. Cambridge, MA: MIT Press.

Gutierrez L. (2003). Panel unit root tests for cross-sectionally correlated panels: A Monte Carlo comparison, University of Sassari, mimeo.

Im, K.S., Pesaran, M.H. and Shin, Y. (1997). Testing for Unit Roots in Heterogeneous Panels. Department of Applied Economics, University of Cambridge.

Levin, A. and Lin, C.F. (1992). Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. Discussion Paper Series 92-23, Department of Economics, University of San Diego.

Levin, A. and Lin, C.F. (1993). Unit Root Tests in Panel Data: New Results. Discussion Paper Series 93-56, Department of Economics, University of San Diego.

Maddala, G.S. and Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel data and a New Simple Test, Oxford Bulletin of Economics and Statistics, Special Issue, 631-652.

Moon, H. R., Perron, P. (2003). Testing for Unit Root in Panels with Dynamic Factors. Research Papers Series, University of Southern California Center for Law, Economics & Organization, n. C01-26.

Phillips, P.C.B., Sul, D. (2003). The elusive empirical shadow of growth convergence. Cowles Foundation Discussion Paper n. 1398.

Quah, D. (1990). International Patterns of Growth: I. Persistence in Cross-Country Disparities. MIT Working paper, January.

Quah, D. (1993a). Galton's fallacy and tests of convergence hypothesis. *The Scandinavian Journal of Econometrics*, 95, 9-19.

Quah, D. (1993b). Empirical cross section dynamics in economic growth. *European Economic Review*, 37, 426-434.

Quah, D. (1994). Exploiting Cross-Section Variations for the Unit Root Inference in Dynamic Data. *Economics Letters*, 44: 9-19.

Quah, D. (1996). Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journals*, 106,1045-1055.

Quah, D. (1997). Empirics for growth and distribution : stratification, polarization, and convergence clubs, *Journal of Economic Growth*, 2, 27-59.

Romer, P. (1990). Endogenous technical change. *Journal of Political Economy*, 98, 71-102.

Silverman, B.W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall, London.

Solow, R. (1956). A Contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1): 312-320.

Swan, T. (1956). Economic growth and capital accumulation. *Economic Record*, 32(November): 344-361.

Table 1 Log EU regions per-capita income trend coefficients by range, 1980-1999

EU Region Range		Trend	*t*_statistics	R^2
First	Last	Coefficient		
1	140	0,057	39,903	0,754
1	70	0,060	31,330	0,837
71	140	0,053	32,930	0,880
1	46	0,062	26,513	0,874
47	92	0,056	29,239	0,964
93	140	0,053	27,541	0,889
1	35	0,063	23,796	0,899
36	70	0,057	25,278	0,955
71	105	0,054	25,299	0,956
106	140	0,052	23,249	0,879
1	28	0,064	21,684	0,916
29	56	0,058	22,534	0,952
57	84	0,056	22,994	0,972
85	112	0,053	22,516	0,951
113	140	0,053	20,864	0,882
1	23	0,066	19,739	0,920
24	46	0,058	20,393	0,951
47	69	0,057	20,727	0,966
70	92	0,054	20,801	0,970
93	115	0,053	20,501	0,956
116	140	0,053	19,985	0,894
1	20	0,067	18,429	0,921
21	40	0,058	19,061	0,953
41	60	0,057	19,275	0,964
61	80	0,057	19,485	0,974
81	100	0,052	19,018	0,951
101	120	0,053	18,560	0,928
121	140	0,052	18,117	0,906
1	17	0,067	17,125	0,929
18	34	0,059	17,619	0,956
35	51	0,058	17,630	0,956
52	68	0,056	17,899	0,971
69	85	0,056	17,966	0,974
86	102	0,052	17,598	0,954
103	119	0,054	17,088	0,927
120	140	0,051	18,248	0,890

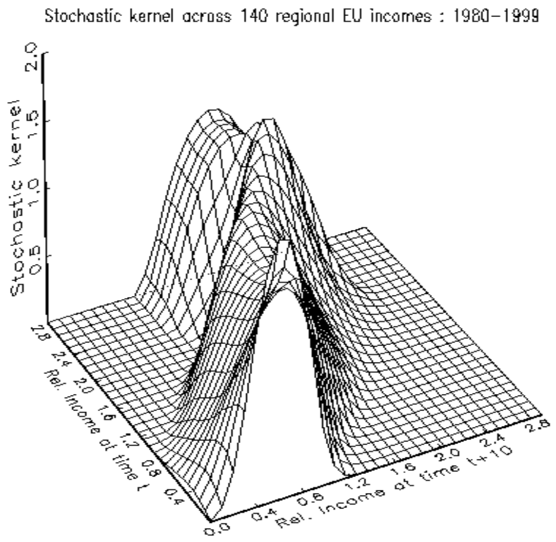

Fig. 1 Stochastic kernel

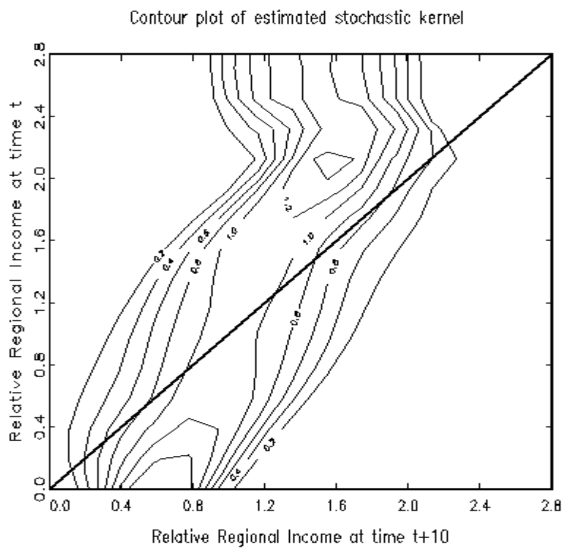

Fig. 2 Contour plot stochastic kernel

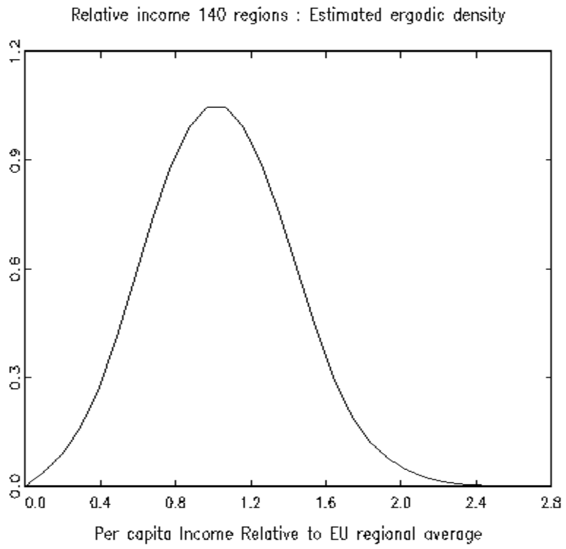

Fig.3 Ergodic (long-run) density distribution

Table 2 Panel Unit Root Tests EU regional Income: 1980-1999	
Group	Moon and Perron (2003) **t_b** test(*)
All (140) regions	-3.372 (0.000)
Sub-groupings according to income:	
Core regions	- 2.181 (0.015)
Peripheral regions	-3.505 (0.000)
Old regions	-1.810 (0.035)
New comers	-2.772 (0.003)